

Subject-Specific Abnormal Region Detection in Traumatic Brain Injury Using Sparse Model Selection on High Dimensional Diffusion Data

Matineh Shaker^{a,b,*}, Deniz Erdogmus^a, Jennifer Dy^a, Sylvain Bouix^b

^a*Electrical and Computer Engineering Department, Northeastern University, Boston, MA. shaker@ece.neu.edu, erdogmus@ece.neu.edu, jdy@ece.neu.edu*

^b*Brigham and Women's Hospital, Harvard Medical School, Boston, MA. sylvain@bwh.harvard.edu*

Abstract

We present a method to estimate a multivariate Gaussian distribution of diffusion tensor features in a set of brain regions based on a small sample of healthy individuals, and use this distribution to identify imaging abnormalities in subjects with mild traumatic brain injury. The multivariate model receives *a priori* knowledge in the form of a neighborhood graph imposed on the precision matrix, which models brain region interactions, and an additional L_1 sparsity constraint. The model is then estimated using the graphical LASSO algorithm and the Mahalanobis distance of healthy and TBI subjects to the distribution mean is used to evaluate the discriminatory power of the model. Our experiments show that the addition of the *a priori* neighborhood graph results in significant improvements in classification performance compared to a model which does not take into account the brain region interactions or one which uses a fully connected prior graph. In addition, we describe a method, using our model, to detect the regions that contribute the most to the overall abnormality of the DTI profile of a subject's brain.

Keywords: sparse learning, graphical lasso, TBI, DTI

*Corresponding author

1. Introduction

Abnormalities of Diffusion Tensor Imaging (DTI) data in neuroimaging studies are traditionally detected at the population level by directly comparing regions of interest across patients and healthy controls, and verifying whether distributions are statistically different in these regions. The assumption behind these types of analyses is that conditions in patients have homogeneous spatial patterns of abnormalities. However, in diseases such as traumatic brain injury (TBI) or multiple sclerosis, a common spatial pattern of injury is unlikely to occur, violating the main hypothesis of standard population studies.

With an estimated 10 million people world-wide affected annually by a TBI, the burden that this condition imposes on society makes it a considerable public health problem Hyder et al. (2007); Feigin et al. (2013); Marion et al. (2011). Importantly, a significant percentage (10-15%) of individuals diagnosed with mild TBI experience persistent post-concussive symptoms (PPCS), which may lead to long-term disabilities Bigler (2008). Symptoms range from physical, such as headache; cognitive, such as difficulty concentrating; and emotional/behavioral, such as irritability and impulsivity. In the majority of these chronic cases, there is no radiological evidence of injury from conventional magnetic resonance imaging (MRI) or computed tomography (CT), and little is known about the pathophysiology underlying the injury. Thus establishing radiological evidence of brain injury is a critical first step towards the proper diagnosis and monitoring of TBI, and may lead to establishing neuroimaging biomarkers to help predict recovery versus PPCS and to assess better the impact of therapies on the injured brain.

Recent methods for injury detection in mild TBI patients have been developed by estimating a model of "healthy" DTI features and testing whether brain regions have outside-of-normal-range values for a particular subject's brain (see Mayer et al. (2014) for a nice overview). Typically, each region is modeled by the mean and standard deviation of the DTI feature of interest over all healthy individuals, and individual TBI subject's data are z-transformed using these healthy population parameters. Finally regions with a z-score above a given threshold (typically 2 standard deviations) are flagged as abnormal, and statistics such as the number of abnormal regions or the average z-score over the brain are compared between TBI and controls. Methods mostly differ from each other based on how the mean and standard deviation are estimated, and how bias is avoided when testing normal con-

trols that have been used to estimate the "healthy" model parameters Ge et al. (2005); Kim et al. (2013); Bouix et al. (2013). Most methods study one DTI feature at a time (except for Hellyer et al. (2013), which uses four DTI features in a multivariate setting), but none of the current techniques model the inter-dependence of DTI features between neighboring brain regions. Another interesting result from our previous work, suggest that DTI changes are observable in gray matter regions in these patients (potentially related to glial scarring), and thus one should study the full brain as opposed to only white matter in this population Bouix et al. (2013).

In this paper, we extend the multiple univariate setting of Bouix et al. (2013) to a high dimensional Gaussian multivariate model which accounts for inter-region interactions. One of the main challenge we need to overcome is a relatively small number of healthy subjects (in the order of 50) compared to the number of parameters to estimate (in the order of 10,000). Our method thus relies on the estimation of a sparse representation of the region co-dependencies as modeled by a precision matrix.

Although not as thoroughly studied in diffusion MRI, sparse representation of inter-region interactions is the subject of much research in fMRI. Extracted networks capture higher order dependencies among variables, and therefore are effective in exploring local interactions of brain regions Friston (2011). Unfortunately, the estimation of these functional connectivities from subject to subject can be difficult to do robustly and recent research has focused on imposing a prior to the sparse representation. One such example is the work of Zhu et al. (2013), which uses structurally-weighted least absolute shrinkage and selection operator (LASSO) regression, and models the directional functional interactions of resting state fMRI data based on *structural* connectivity constraints encoded by 358 cortical landmarks derived from DTI data Zhu et al. (2012).

Our work is similar in spirit, with some key differences. Here, we use DTI to evaluate subtle tissue changes in TBI patients by detection of outliers compared to a model of normal brain derived from 145 brain regions of 34 healthy subjects. A feature vector containing fractional anisotropy (FA) measures over 145 brain regions represents each subject. We model the distribution of these features in the healthy subjects as a multi-dimensional Gaussian distribution as represented by a precision matrix. Our method relies on the theorem that conditional independence of two variables given others is equivalent to setting the corresponding precision matrix entity to zero Lauritzen (1996). We leverage this theorem by imposing a brain neigh-

neighborhood prior graph on the structure of the precision matrix, reducing the number of parameters to estimate by favoring interactions of proximal regions and ignoring the interactions of regions which are far away from each other.¹ The multi-dimensional Gaussian model is further regularized by an L_1 sparsity constraint and estimated using the graphical LASSO Friedman et al. (2008).

2. Gaussian graphical models

Let $\mathbf{x} = [X_1, X_2, \dots, X_d]$ be a d -dimensional random vector so that it has a multivariate Gaussian distribution $\mathbf{x} \sim \mathcal{N}(\mu, \Sigma)$, with d -dimensional mean vector μ , and a $d \times d$ covariance matrix Σ . In a Gaussian graphical model, an unweighted undirected graph with adjacency matrix G , can be used to represent the conditional dependence structure between the individual variables X_i . More specifically, the edge structure of G can be imposed onto the inverse covariance matrix, also known as the precision matrix, $\Sigma^{-1} \equiv \Theta = \{\theta_{ij}\}$, and *conditional independence* between X_i and X_j can be expressed as a zero in the corresponding location in Θ :

$$X_i \perp\!\!\!\perp X_j \Leftrightarrow G_{ij} = 0 \Leftrightarrow \theta_{ij} = 0 \quad (1)$$

The proof can be found in Lauritzen (1996).

One key benefit of this representation is that one can use a priori information to impose a conditional independence structure to the model. This is particularly useful in scenarios where a high dimensional Θ needs to be estimated with only a few samples, and expert knowledge about the data set can help guide sparse model learning. By using a graph G which sets many of the precision matrix elements to zero before the estimation process, we can greatly reduce the number of parameters of the model, and thereby increase the robustness of the optimization.

In addition, we assume global sparsity of the model, and thus add an L_1 penalty term to further regularize the model. Following Banerjee et al. (2008), let X be the $n \times d$ data matrix representing n observations, S be the $d \times d$ sample covariance matrix, and G the a priori graph, the maximum

¹Note that a sparse precision matrix does not imply a sparse covariance matrix; therefore distant brain regions are not assumed to be independent with this constraint – only conditionally independent as shown in Eq. (1).

a-posteriori (MAP) estimate of Θ given X and G is:

$$\begin{aligned}\hat{\Theta}_{MAP} &= \arg \max_{\Theta} \log p(\Theta|X, G) \\ &= \log \det \Theta - \text{tr}(S\Theta) - \rho \|\Theta\|_{L_1} \\ &\quad \text{with } \theta_{ij} = 0 \text{ when } G_{ij} = 0.\end{aligned}\tag{2}$$

where ρ is a scalar controlling the L_1 norm penalty weight.

The optimization method uses the graphical LASSO algorithm Friedman et al. (2008), which can elegantly incorporate G into the optimization process.

In the following section, we describe how this graphical model with the addition of an a priori graph can be applied to the problem of estimating a multivariate Gaussian distribution of DTI features in healthy subjects and use this model to detect brain injuries in subjects with mild TBI.

3. Application to injury detection in TBI

Our driving hypothesis for using graphical models is that brain regions next to each other have similar, or at least highly related, DTI signal in healthy subjects. We thus model these interactions by only considering edges connecting proximal regions in the graph imposed on the precision matrix. If a TBI subject has a region with abnormal signal, having modeled the healthy region-to-region interaction will help us increase our sensitivity to classifying a TBI brain as abnormal, compared to looking at each region independently.

3.1. Subjects and data acquisition

In this work, we used the data described in Bouix et al. (2013). There are $n = 34$ healthy subjects, $p = 11$ TBI patients who reported symptoms (see Table 1 for details), such as headaches, emotional dysregulation and memory impairments at the time of data collection, as well as $m = 11$ normal controls demographically matched to TBIs. The normal controls are separated from healthy subjects for validation purposes. Subjects underwent MRI scanning, including a high resolution diffusion tensor imaging scan and a high resolution structural T1 weighted scan. Each T1 image was segmented using the FreeSurfer software Fischl et al. (2002), resulting in 176 gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF) sections. CSF sections and sections smaller than 300 mm^3 were excluded from the analysis, as these

Table 1: Description of individual mTBI subjects adapted from Bouix et al. (2013)

ID	Age	Gender	Source of Injury	Duration since Injury	Symptoms
TB01	45	F	MVA*	17.0	Cognitive impairment, emotional dysregulation, depression
TB02	38	M	MVA	106.6	Mild memory impairment, mild executive function impairment emotional dysregulation
TB03	44	F	MVA	121.3	Dizziness, exhaustion, periodic limb movements, hypersomnia depression and anxiety
TB04	30	M	Sports Injury	2.6	Diplopia, fatigues easily, executive function impairment
TB05	42	M	MVA	138.0	Cognitive impairment, memory executive function impairment
TB06	28	M	Assault	27.0	Anxiety, depression, insomnia ADHD, intrusive thoughts, memory deficits, overeating
TB07	24	M	Blast Exposure	70.3	Anxiety, dpanic attacks, hypervigilance, overeating, difficulty concentrating
TB08	25	M	Blast Exposure	83.3	Depression, memory impairment difficulty w/rapidly presented information
TB09	29	M	Blast Exposure	51.4	Irritability, nightmares, depression, panic attacks, cognitive and memory impairment
TB10	24	M	Blast Exposure	55.9	Headaches, memory impairment, problems concentrating, irritability, anxiety, nightmares
TB11	39	M	Sports Injury	9.5	Facial pain, memory/executive function impaired, emotional dysregulation

*MVA= Motor Vehicle Accident

smaller regions led to unstable estimation of mean/std of the DTI measures and many failed to pass normality tests. The remaining 145 sections (83 in GM and 62 in WM) were registered onto the diffusion space using a non-linear diffeomorphic registration algorithm Avants et al. (2011). The average FA was computed in each region for each subject. The outcome of the image processing procedure is a feature vector of the average FA in $d = 145$ brain structures in each subject. More details about data acquisition and processing can be found in Bouix et al. (2013). In addition, the same procedure was applied for the other standard DTI measures: mean diffusivity (MD), radial diffusivity (RD), and axial diffusivity (AD).

3.2. *a priori* graph

Given this data set, we will have to estimate a 145×145 precision matrix based on 34 observations. In order to reduce the number of parameters to estimate, we chose to design a simple graph that will only consider the relationship between neighboring regions in the brain. Two regions were considered to be neighbors if they were connected in a template FreeSurfer segmentation using 26-connectivity. Our motivation for choosing this graph for TBI stems from the knowledge that nearby regions in healthy subjects will tend to have similar tissue properties and thus similar DTI signal (note that we are not considering tensor orientation).

We have made the choice of connecting neighboring GM and WM regions with an edge as there is increasing evidence that the tissue and geometric properties of proximal GM/WM regions are strongly related Miyata et al. (2009); Koch et al. (2013); Liu et al. (2014); Savadjiev et al. (2014). Furthermore, the graph is only a guide for the precision matrix estimation process. If the data does not support the existence of a (conditional) relationship between two variables, the corresponding entry in the precision matrix will converge to zero even if it was linked by an edge in the prior graph.

The neighborhood network G is illustrated in Figure 1. Each brain structure is represented as a node in the graph, and conditional dependence is only considered between regions connected by an edge, whereas all other relationships are ignored. Bold lines in Figure 1(b) show the subgraph associated with region 1. The adjacency matrix corresponding to the complete neighborhood graph is shown in Figure 1(c). One can observe a large number of parameters that will be set to 0 in Θ . Note that the conditional independence of two non-neighboring regions imposed by this graph does not enforce unconditional independence; pairs of regions that are not immediate neighbors are allowed to have correlations.

3.3. *Identifying an abnormal brain*

Let X be the $n \times d$ matrix representing the set of d features in n healthy subjects, Y the $m \times d$ matrix capturing the observations in m normal controls, and Z the $p \times d$ matrix representing the set of p TBI patients. Normal controls are healthy subjects matched to patients demographically, and are separated from the healthy training set X for validation purposes.

The overall design is to generate a model (μ_X, Θ_X) based on the healthy subject data X and test whether a TBI subject i is abnormal by measuring

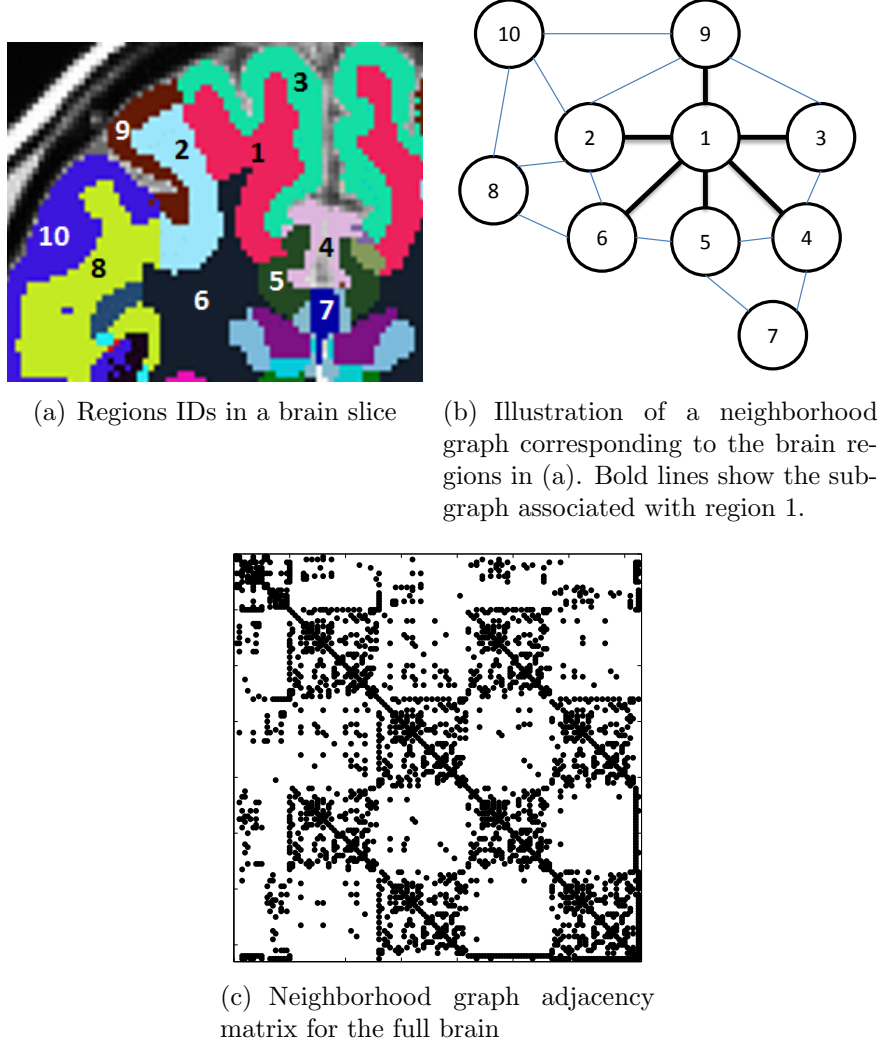


Figure 1: Illustration of the prior graph through 10 brain regions. The vector assigned to region 1 in the adjacency matrix is $[1110110000]$, emphasizing connections of regions 2, 3, 5, 6 and disconnections of regions 4, 7, 8, 9, 10 to region 1. Bold lines in Figure 1(b) show the subgraph associated with region 1.

the Mahanobis distance of its feature vector \mathbf{z}_i to the model:

$$d_M(\mathbf{z}_i) = \sqrt{(\mathbf{z}_i - \mu_X)^T \Theta_X (\mathbf{z}_i - \mu_X)} \quad (3)$$

As the Mahalanobis distance follows a χ^2 distribution, a threshold for

an abnormal brain based on this distance can be theoretically derived (e.g., above the 95th percentile of the expected Mahalanobis distances). However, in our work, we test the discriminatory power of our model by computing the Mahalanobis distances of TBI subjects (Z) and matched controls (Y) and evaluate its classification performance using Receiver Operating Characteristic (ROC) curve analysis.

3.4. Identifying individual abnormal regions

The method we have presented thus far has the ability to identify whether a subject's imaging profile is *overall* abnormal. The natural next step is to identify which regions are most affected in this subject and thus provide some information that could potentially be linked to the pathophysiology of the brain injury, or help targeting therapies to particular brain areas. Given k regions, we propose a greedy forward sorting approach to identify these abnormal regions as follows. Let R_s be the ordered set of sorted regions from most normal to most abnormal, $R_u = \{1, \dots, k\}$ be the set of all regions, and d_R be the Mahalanobis distance computed by only taking into account the regions in subset $R \subset R_u$. We build R_s by incrementally adding the region $r_i \in R_u \setminus R_s$, which minimizes d_{R_i} , where $R_i = R_s \cup \{r_i\}$. This process is repeated until all regions have been sorted from most normal to most abnormal. The procedure is detailed in Alg. 1

Algorithm 1 Sorting regions from most normal to most abnormal

```

1:  $R_u = \{1, \dots, k\}$ 
2:  $R_s = ()$ 
3: for  $i$ : 1 to  $k$  do
4:    $r_i = \arg \min_{j \in R_u \setminus R_s} (d_{R_s \cup \{j\}})$ 
5:    $R_s = R_s \cup \{r_i\}$ 
6: end for
7: return  $R_s$ 

```

The output of this algorithm is an ordering of regions along with k Mahalanobis distances, d_{R_i} of the corresponding subsets of sorted regions. The last step consists of comparing the subject's sorted D_i s with the theoretical distribution of the Mahalanobis distance (the χ^2 distribution with i degrees of freedom) and finding the first region after which the subject's sorted distances exceed the 95th percentile of the χ^2 distribution. Let $F_{\chi^2}(D, l)$ be the cumulative distribution function of the χ^2 distribution with l degrees

of freedom and $\hat{k} = \arg \max_k (F_{\chi^2}(D_k, k) < 0.95)$. Thanks to our sorting process, the regions that are not in the subset of size \hat{k} will generate increasingly unlikely Mahalanobis distances and can be flagged as abnormal. This thresholding procedure is illustrated in Figure 2.

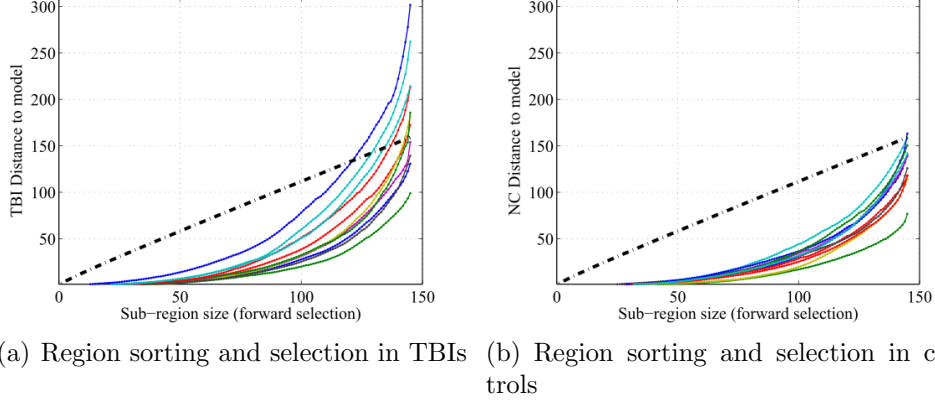


Figure 2: Abnormal Region detection based on the χ^2 distribution and our greedy forward sorting of regions. The dashed lines correspond to the 95th percentile of the CDF of the χ^2 distribution. Each of the colored curves represent, for each subject, the accumulated Mahalanobis distance given by the subsets of sorted regions. Regions added after the point at which the subjects Mahalanobis distance curve exceeds the 95% threshold on the CDF of χ^2 distribution are flagged as abnormal.

4. Experiments

In order to evaluate the performance of the prior neighborhood graph approach, we tested three different graph structures as follows:

1. The neighborhood prior graph as described in Section 3.2, with an L_1 sparsity constraint.
2. A node-only graph with all off-diagonal elements set to zero in the precision matrix.
3. A fully connected graph evaluating all off-diagonal terms with an L_1 sparsity constraint.

We tested the robustness of each model by performing a cross validation procedure as follows. Using a leave-one-out strategy, we generated $n - 1$ models $(\mu_i, \tilde{\Theta}_i)$ from $X_{|i}$, the set of healthy subjects X without the i -th

element. For each model $(\mu_i, \tilde{\Theta}_i)$, we then calculate $d_{M, X_{|i}}$ for all TBI subjects in Z and for all control subjects in Y . In addition, we repeated this procedure for a range of regularization parameter ρ (from 10^{-2} to 10) to evaluate the impact of this parameter on performance. Thus, for each ρ we had n sets of “TBI vs. Controls” Mahalanobis distances and were able to compute confidence intervals of various classification performance measures (in our case the area under the receiver operating characteristic curve – AUC).

As described earlier, the maximization of the posterior distribution in (2), iteratively minimizes certain edges of the graph in two ways: 1) Data driven, where natural interaction of variables among all samples estimate the edges in the graph or precision matrix elements; 2) Prior model driven, where a predefined graph is imposed to the model which sets certain edges to zero, without iterative learning.

In the following experiments, the performance of the node only graph (diagonal precision matrix) is evaluated to illustrate the importance of multivariate vs. univariate analyses. Graphical LASSO is clearly not needed in this diagonal precision matrix design.

4.1. Node-only versus neighborhood versus fully-connected graphs

In Figure 3, all three graph types are examined. In addition, the evaluation is performed for different ρ values to observe the impact of this regularization parameter on the classifier performance.

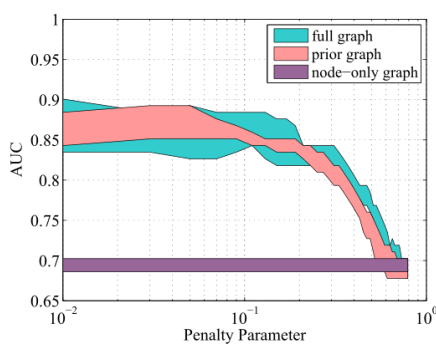
Figure 3(a) compares the 90% confidence intervals (CI) of the $AUC(\rho)$ functions of 34 cross-validation instances across graph types. The confidence intervals are computed using the functional box plot method Sun and Genton (2011), and the envelope of the 90% central region is shown in Figure 3. One can observe that both the neighborhood and the full graph clearly outperform the node-only model. In addition, while these two graphs have comparable average performance over all cross-validation, the neighborhood graph has a tighter 90% confidence interval.

The advantage of the prior graph over a fully-connected graph is even clearer when considering the Bayesian information criterion (BIC), as given by

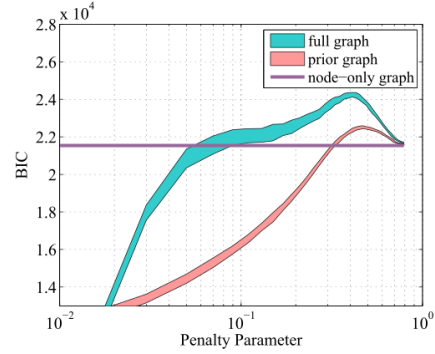
$$BIC = -2 \ln p(X|\hat{\Theta}) + j \ln n, \quad (4)$$

where $p(X|\hat{\Theta})$ is the maximized value of the likelihood function, j is the number of parameters estimated, and n is the number of training samples. In our case, j represents the number of non-zero values in the estimated

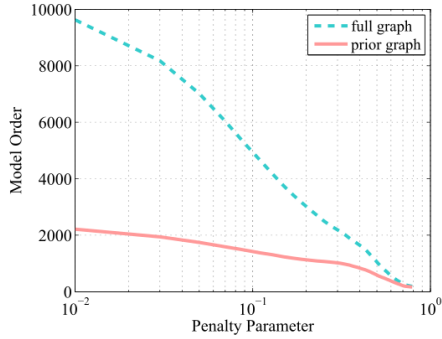
precision matrix $\hat{\Theta}$. BIC is a criterion for model selection among a finite set of models, and balances the goodness of fit ($p(X|\hat{\Theta})$) with a penalty term for the number of model parameters. This criterion penalizes models which increase their likelihood by overfitting the data. Using BIC as a model selection criteria, the prior graph model is preferred due to its lower BIC. Figure 3(c) compares the number of parameters estimated (model order) for the two multivariate models. In Figure 3(d), one can observe that the neighborhood graph always has a higher AUC than the full graph for the same model complexity.



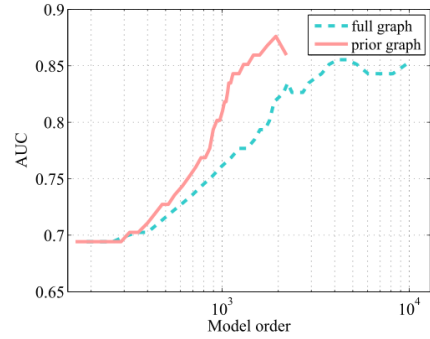
(a) AUC as a function of ρ



(b) BIC as a function of ρ



(c) Complexity of models



(d) Performance vs. model complexity

Figure 3: (a) 90% CIs of the AUC, as a function of the penalty parameter ρ for the prior graph, fully-connected graph, and node-only graph; (b) 90% CIs of the BIC as a function of the penalty parameter ρ for each model; (c) Number of model parameters for a given penalty weight for prior and fully-connected graphs; (d) AUC as a function of the number of model parameters for the prior and fully-connected graphs. Both models show similar AUC performance, but the prior graph has better model complexity and BIC.

4.2. Neighborhood versus random prior graphs

To check the importance of expert knowledge in model selection, 1000 random graphs were generated so that they have the same number of edges as the prior graph but at uniformly random locations. Figure 4 compares the AUC and BIC of the neighborhood graph to the 75%, 85% and 95% central regions of the randomly generated graphs. The neighborhood graph has an AUC that is higher than the 95% central region of random graphs. Similarly the BIC is almost always lower for the neighborhood graph than it is for random graphs. This result illustrates that the better performance of the neighborhood prior is not due to overfitting, but because of the selection of an appropriate graphical model. The percentile of the prior graph performance at various ρ compared to the random graphs distribution is shown in Table (2).

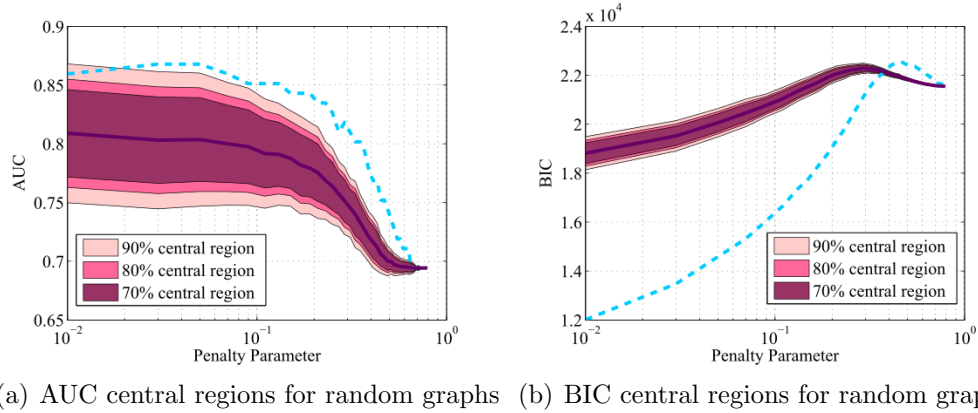


Figure 4: Comparison of the neighborhood graph model performance (blue dashed line) with 1000 random graphs with the same number of edges.

Table 2: percentile of the prior graph performance on the random graph distribution

ρ	0.01	0.02	0.1	0.2	0.4	0.8
Percentile	98%	99%	100%	98%	100%	100%

4.3. Selecting the optimal penalty parameter ρ

While the above analyses provide valuable information on the quality of the different models under different regularization by the parameter ρ , one

does need to select a single optimal $\hat{\rho}$ value to estimate the final model. In order to find this optimum, the Mahalanobis distance of each training point to the model mean estimated with the remaining training points is calculated. The optimum ρ minimizes the leave-one-out sum of squared distances, which is $\hat{\rho} = 0.3$ for the prior graph and $\hat{\rho} = 0.38$ for the full graph, as shown in Figure 5. Table 3 compares the performance of the three models at optimum values of ρ . Once again, the neighborhood prior graph model outperforms both the node-only and full graph priors. Note that the performance of the node-only graph does not depend on the value of ρ .

In order to put our results in context with traditional "z-score" approaches White et al. (2009); Lipton et al. (2012); Bouix et al. (2013); Mayer et al. (2014), we also performed the computation of the AUC of the mean absolute z-score over all regions as a potential measure to distinguish patients from controls.

Z-scores were computed with respect to the mean and standard deviation of FA in each region over X , the training set of healthy subjects. As expected, this method does not perform as well as the multivariate models.

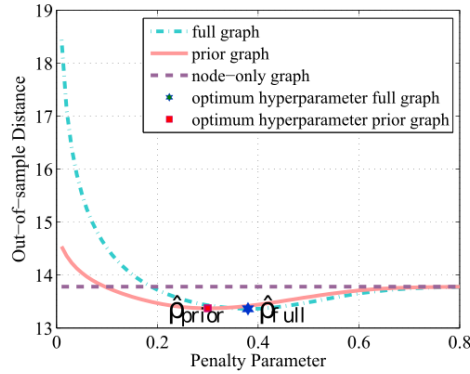


Figure 5: Selection of the optimum penalty parameter for each model based on minimizing leave-one-out sum of squared Mahalanobis distances to the model mean.

Table 3: AUC analysis for optimal multivariate models and independent z-scores

	Full Graph	Prior Graph	Node-only Graph	mean absolute zscore
AUC	0.83	0.86	0.69	0.65
Sensitivity	0.64	0.73	0.73	0.64
Specificity	0.91	1	0.64	0.64

4.4. Investigating other DTI measures

The z-score analysis of Bouix et al. (2013) only found statistically significant differences for FA. Nevertheless, we further tested, using our multivariate method, the other most common DTI measures: Mean Diffusivity (MD), Axial Diffusivity (AD), and Radial Diffusivity (RD). For all experiments, we used the prior graph and the same regularization parameter $\hat{\rho} = 0.3$. As in the previous work, only FA reached significance, although we hypothesize that AD could reach significance given a larger sample size (see Table 4). Consequently, all subsequent analyses focused solely on FA.

Table 4: p-value of Wilcoxon ranksum tests and AUC for the most common DTI metrics.

Measure	p	AUC
FA	0.016	0.86
MD	0.168	0.68
AD	0.088	0.72
RD	0.265	0.64

4.5. Correlations with behavioral measures

Similarly to Bouix et al. (2013), we performed Spearman correlations between the Mahalanobis distance and behavioral measures in BI subjects. The results presented in Table 5 are very similar to our previous work, with "Digit Symbol", a measure of processing speed, the only behavioral test significantly correlated with imaging ($\rho = -0.62$, $p = 0.04$), although reported p-values are uncorrected for multiple comparisons. The Bonferroni corrected significance threshold is 0.004

Nevertheless, our sample of 11 TBI subjects is quite small, and we expect better correlations with a larger number of subjects. Confidence intervals on ρ is calculated according to the formula presented in Ruscio (2008).

4.6. Individual abnormal regions identification

In this section, we present the results of the detection of individual abnormal regions as described in section 3.4. Each subfigure in Figure 6 shows a $k \times l$ matrix. The k rows represent the regions and the l columns the individual subjects. The intensity associated with each region in each figure corresponds to its respective amount of "abnormality". We define this abnormality a_i as the following differential

$$a_i = \frac{(D_i - D_{i-1})}{(\tilde{D}_i - \tilde{D}_{i-1})},$$

Table 5: Spearman correlations between Behavioral measures and Mahalanobis distances based on FA in TBI subjects. Note that p-values are uncorrected for multiple comparisons, the Bonferroni corrected significance threshold is 0.004

Test	Subtest	rho	p (uncorrected)	CI_{min}	CI_{max}
California Verbal Learning Test II	Trials 1-5	0.27	0.42	-0.40	0.75
	Short Delay Free Recall	-0.57	0.07	-0.87	0.04
	Short Delay Cued Recall	-0.37	0.26	-0.80	0.30
	Long Delay Free Recall	-0.35	0.29	-0.78	0.31
	Long Delay Cued Recall	-0.23	0.49	-0.73	0.42
Processing Speed	Digit Symbol	-0.62	0.04*	-0.89	-0.04
	Symbol Search	-0.45	0.16	-0.82	0.20
Digit Span	Digit Span	-0.30	0.37	-0.76	0.36
Trail Making	Trail Making A	-0.02	0.96	-0.61	0.59
	Trail Making B	0.37	0.26	-0.30	0.80
Controlled Oral Word Association		-0.18	0.60	-0.70	0.47
STROOP		0.17	0.61	-0.47	0.70

where D_i is the Mahalanobis distance of the sorted subset of size i and \tilde{D}_i is 95% threshold of the χ^2 distribution, i.e., $F_{\chi^2}(\tilde{D}_i, i) = 0.95$.

We also present the equivalent figures for standard z-score analyses in Figure 7. In this figure, we present regions with an absolute z-score greater than 2 as well as those greater than 3.58, the threshold corresponding to a Bonferroni correction for the number of regions.

One can observe that both the neighborhood and full graph display similar patterns of detections, whereas the node-only graph displays many false positives. The z-score method show similar results to the node only graph at $|z| > 2$ and a subset of the multivariate techniques at $|z| > 3.58$.

5. Discussion

Graphical models are a powerful and flexible technique to impose a structure on a multivariate Gaussian model, which has allowed us to constrain the estimation of a model of DTI signal based on a small data set of healthy subjects.

We chose to constrain a LASSO estimation procedure of a precision matrix, by imposing a *conditional independence* structure on our model Lauritzen (1996). Note the emphasis on *conditional* independence, i.e., the lack of an edge in our prior graph does not forbid covariance between two variables, but assumes that for two variables X & Y, knowing X offers no additional information about Y given what we already know from the other variables in the model. Therefore, the independence structure imposed by the graph is

quite flexible and allows for the examination of many relationships including those of regions that are very far apart.

We applied this method to detect whether subjects who experienced a TBI had an abnormal DTI scan, by measuring the Mahalanobis distance of their data to the model. We tested three different graph structures, a node-only graph, a fully connected graph, and a neighborhood graph, which only connects regions that are next to each other in the brain. The ability of each method to accurately detect an abnormal brain was tested by classifying TBI vs NC subjects using their Mahalanobis distance to the model under study and computing the corresponding AUC.

Our results demonstrate that multivariate approaches (full and neighborhood graph) clearly outperform the univariate approaches, including standard z-score analyses White et al. (2009); Lipton et al. (2012); Bouix et al. (2013); Mayer et al. (2014). While both full and neighborhood graph show similar AUCs, the neighborhood graph leads to a better model when taking into account model complexity, i.e., the number of non-zero elements in the precision matrix. Furthermore, our cross-validation experiments show that although the sample size is small, the results are quite robust as the 90% central region width of the AUC is less than 0.05 for the neighborhood graph. Moreover, the neighborhood model always outperforms randomly generated graph with the same number of edges, indicating that the “expert” knowledge embedded in the graph is indeed a valuable prior to constrain the estimation of the model.

Importantly, the flexibility of graphical models can allow us to test a number of prior graphs, including network-based graph generated from diffusion MRI and/or functional MRI network analyses Yoldemir et al. (2015); Vergara et al. (2016). This is certainly a topic we plan to investigate in future work. Another possible extension is the study of DTI (or more generally diffusion MRI) measures in combination, by using a nested precision matrix design, although larger sample sizes would be needed for such complex models. We are particularly interested in diffusion MRI measures related to neuroinflammation such as free water, as it may be a marker for subjects experiencing chronic symptoms Pasternak et al. (2014); Planetta et al. (2016)

We have also shown that our multivariate analysis can detect individual regions with abnormal data. In fact, our results show fewer false positives in NCs and more regions detected in mTBIs compared to classical independent z-score analyses. Nevertheless, this aspect of our work was exploratory and further development inspired by factor analysis techniques should be

investigated.

Finally, we tested the connection between imaging data and symptomatology, but unfortunately were not able to find strong relationships between behavioral measures and DTI beyond a single measure (Digit Symbol, a measure of processing speed). We believe the main reason is the small sample size, but also the fact that we have only looked at the overall Mahalanobis distance (a global imaging measure). With more data, one could investigate connections between symptoms and subsets of regions corresponding to known networks associated with a particular brain function ((e.g., Han et al. (2016)), which we think will lead to stronger relationships between imaging and behavioral measures.

Acknowledgements

This work was supported in part by a CIMIT Soldier in Medicine Award; NSF grants CCF 1442728, IIS-1149570, and IIS-1118061; NIH grants R01 NS 078337 and RO1HL089856; DoD grants W81XWH-08-2-0159; and a Veterans Administration Merit Review Award.

References

- Avants, B.B., Tustison, N.J., Song, G., Cook, P.A., Klein, A., Gee, J.C., 2011. A reproducible evaluation of ants similarity metric performance in brain image registration. *NeuroImage* 54, 2033–2044.
- Banerjee, O., El Ghaoui, L., d’Aspremont, A., 2008. Model selection through sparse maximum likelihood estimation for multivariate gaussian or binary data. *The Journal of Machine Learning Research* 9, 485–516.
- Bigler, E.D., 2008. Neuropsychology and clinical neuroscience of persistent post-concussive syndrome. *J Int Neuropsychol Soc* 14, 1–22. doi:10.1017/S135561770808017X.
- Bouix, S., Pasternak, O., Rathi, Y., Pelavin, P.E., Zafonte, R., Shenton, M.E., 2013. Increased gray matter diffusion anisotropy in patients with persistent post-concussive symptoms following mild traumatic brain injury. *PloS one* 8, e66205.
- Feigin, V.L., Theadom, A., Barker-Collo, S., Starkey, N.J., McPherson, K., Kahan, M., Dowell, A., Brown, P., Parag, V., Kydd, R., Jones, K., Jones,

- A., Ameratunga, S., BIONIC Study Group, 2013. Incidence of traumatic brain injury in New Zealand: a population-based study. *Lancet Neurol* 12, 53–64. doi:10.1016/S1474-4422(12)70262-4.
- Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., Van Der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., et al., 2002. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron* 33, 341–355.
- Friedman, J., Hastie, T., Tibshirani, R., 2008. Sparse inverse covariance estimation with the graphical lasso. *Biostatistics* 9, 432–441.
- Friston, K.J., 2011. Functional and effective connectivity: a review. *Brain connectivity* 1, 13–36.
- Ge, Y., Law, M., Grossman, R.I., 2005. Applications of diffusion tensor mr imaging in multiple sclerosis. *Annals of the New York Academy of Sciences* 1064, 202–219.
- Han, K., Chapman, S.B., Krawczyk, D.C., 2016. Disrupted Intrinsic Connectivity among Default, Dorsal Attention, and Frontoparietal Control Networks in Individuals with Chronic Traumatic Brain Injury. *J Int Neuropsychol Soc* 22, 263–279.
- Hellyer, P.J., Leech, R., Ham, T.E., Bonnelle, V., Sharp, D.J., 2013. Individual prediction of white matter injury following traumatic brain injury. *Annals of neurology* 73, 489–499.
- Hyder, A.A., Wunderlich, C.A., Puvanachandra, P., Gururaj, G., Kobusingye, O.C., 2007. The impact of traumatic brain injuries: a global perspective. *NeuroRehabilitation* 22, 341–353.
- Kim, N., Branch, C.A., Kim, M., Lipton, M.L., 2013. Whole brain approaches for identification of microstructural abnormalities in individual patients: comparison of techniques applied to mild traumatic brain injury. *PloS one* 8, e59382.
- Koch, K., Schultz, C.C., Wagner, G., Schachtzabel, C., Reichenbach, J.R., Sauer, H., Schlosser, R.G., 2013. Disrupted white matter connectivity is associated with reduced cortical thickness in the cingulate cortex in schizophrenia. *Cortex* 49, 722–729.

- Lauritzen, S.L., 1996. Graphical models. Oxford University Press.
- Lipton, M.L., Kim, N., Park, Y.K., Hulkower, M.B., Gardin, T.M., Shifteh, K., Kim, M., Zimmerman, M.E., Lipton, R.B., Branch, C.A., 2012. Robust detection of traumatic axonal injury in individual mild traumatic brain injury patients: intersubject variation, change over time and bidirectional changes in anisotropy. *Brain imaging and behavior* 6, 329–342.
- Liu, X., Lai, Y., Wang, X., Hao, C., Chen, L., Zhou, Z., Yu, X., Hong, N., 2014. A combined DTI and structural MRI study in medicated-naïve chronic schizophrenia. *Magn Reson Imaging* 32, 1–8.
- Marion, D.W., Curley, K.C., Schwab, K., Hicks, and the mTBI Diagnostics Work, R.R., 2011. Proceedings of the Military mTBI Diagnostics Workshop, St. Pete Beach, August 2010. *Journal of Neurotrauma* 28, 517–526. URL: <http://www.liebertonline.com/doi/abs/10.1089/neu.2010.1638>, doi:10.1089/neu.2010.1638.
- Mayer, A.R., Bedrick, E.J., Ling, J.M., Toulouse, T., Dodd, A., 2014. Methods for identifying subject-specific abnormalities in neuroimaging data. *Human brain mapping* 35, 5457–5470.
- Miyata, J., Hirao, K., Namiki, C., Fujiwara, H., Shimizu, M., Fukuyama, H., Sawamoto, N., Hayashi, T., Murai, T., 2009. Reduced white matter integrity correlated with cortico-subcortical gray matter deficits in schizophrenia. *Schizophr. Res.* 111, 78–85.
- Pasternak, O., Koerte, I.K., Bouix, S., Fredman, E., Sasaki, T., Mayinger, M., Helmer, K.G., Johnson, A.M., Holmes, J.D., Forwell, L.A., Skopelja, E.N., Shenton, M.E., Echlin, P.S., 2014. Hockey Concussion Education Project, Part 2. Microstructural white matter alterations in acutely concussed ice hockey players: a longitudinal free-water MRI study. *J. Neurosurg.* 120, 873–881.
- Planetta, P.J., Ofori, E., Pasternak, O., Burciu, R.G., Shukla, P., DeSimone, J.C., Okun, M.S., McFarland, N.R., Vaillancourt, D.E., 2016. Free-water imaging in Parkinson’s disease and atypical parkinsonism. *Brain* 139, 495–508.

- Ruscio, J., 2008. Constructing confidence intervals for spearmans rank correlation with ordinal data: A simulation study comparing analytic and bootstrap methods. *Journal of Modern Applied Statistical Methods* 7, 7.
- Savadjiev, P., Rathi, Y., Bouix, S., Smith, A.R., Schultz, R.T., Verma, R., Westin, C.F., 2014. Fusion of white and gray matter geometry: a framework for investigating brain development. *Med Image Anal* 18, 1349–1360.
- Sun, Y., Genton, M.G., 2011. Functional boxplots. *Journal of Computational and Graphical Statistics* 20.
- Vergara, V.M., Mayer, A., Damaraju, E., Kiehl, K., Calhoun, V.D., 2016. Detection of Mild Traumatic Brain Injury by Machine Learning Classification using Resting State Functional Network Connectivity and Fractional Anisotropy. *J. Neurotrauma* .
- White, T., Schmidt, M., Karatekin, C., 2009. White matter potholes in early-onset schizophrenia: a new approach to evaluate white matter microstructure using diffusion tensor imaging. *Psychiatry Research: Neuroimaging* 174, 110–115.
- Yoldemir, B., Ng, B., Abugharbieh, R., 2015. Coupled Stable Overlapping Replicator Dynamics for Multimodal Brain Subnetwork Identification. *Inf Process Med Imaging* 24, 770–781.
- Zhu, D., Li, K., Guo, L., Jiang, X., Zhang, T., Zhang, D., Chen, H., Deng, F., Faraco, C., Jin, C., et al., 2012. Dicccol: dense individualized and common connectivity-based cortical landmarks. *Cerebral cortex* , bhs072.
- Zhu, D., Li, X., Jiang, X., Chen, H., Shen, D., Liu, T., 2013. Exploring high-order functional interactions via structurally-weighted lasso models, in: *Information Processing in Medical Imaging*, Springer. pp. 13–24.

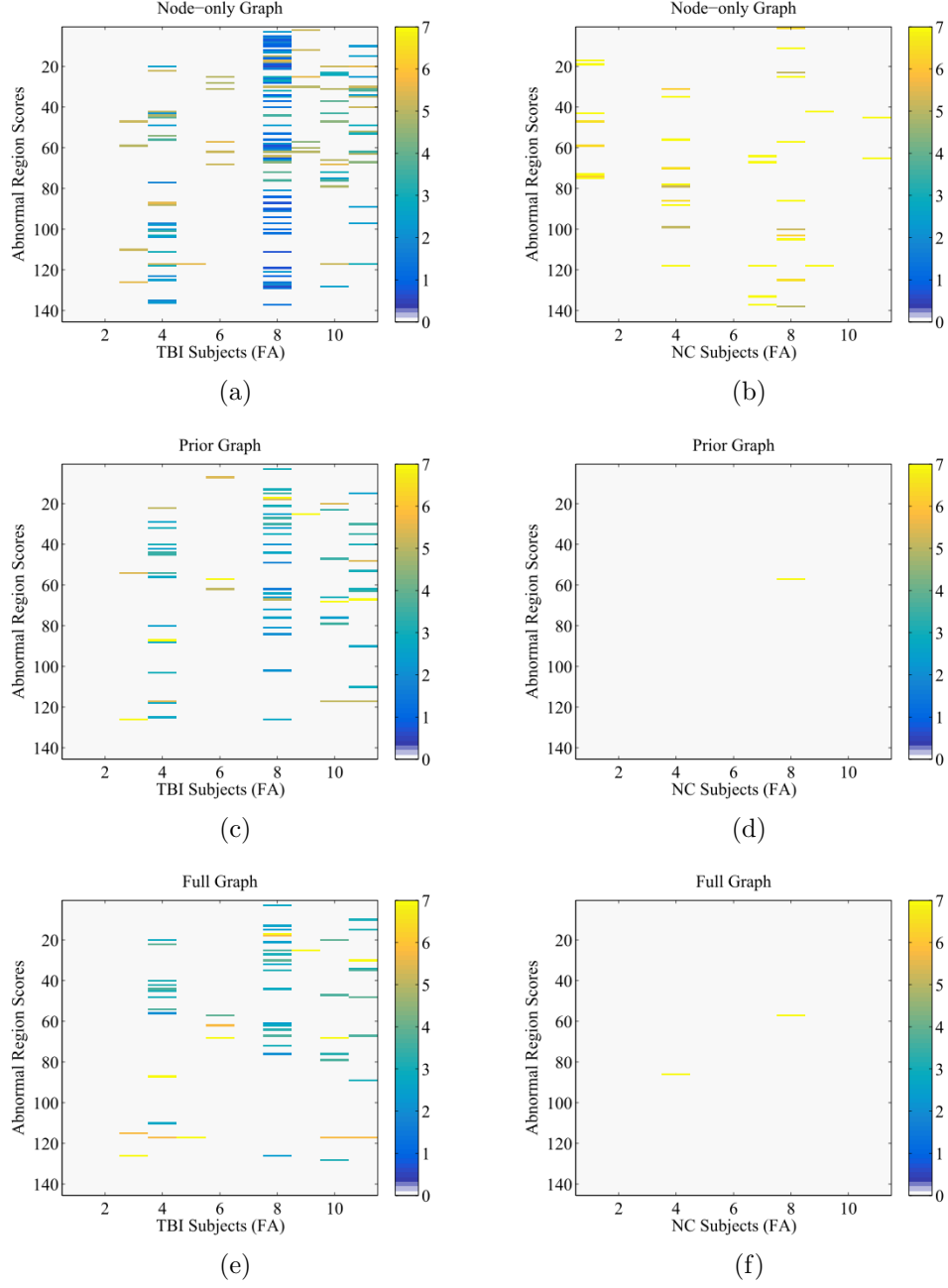


Figure 6: Abnormality maps show which regions are affected in each subject for different graphical models. Left: TBI subjects, Right: Normal Controls (NC). Top to Bottom: node-only graph, neighborhood graph, full graph.

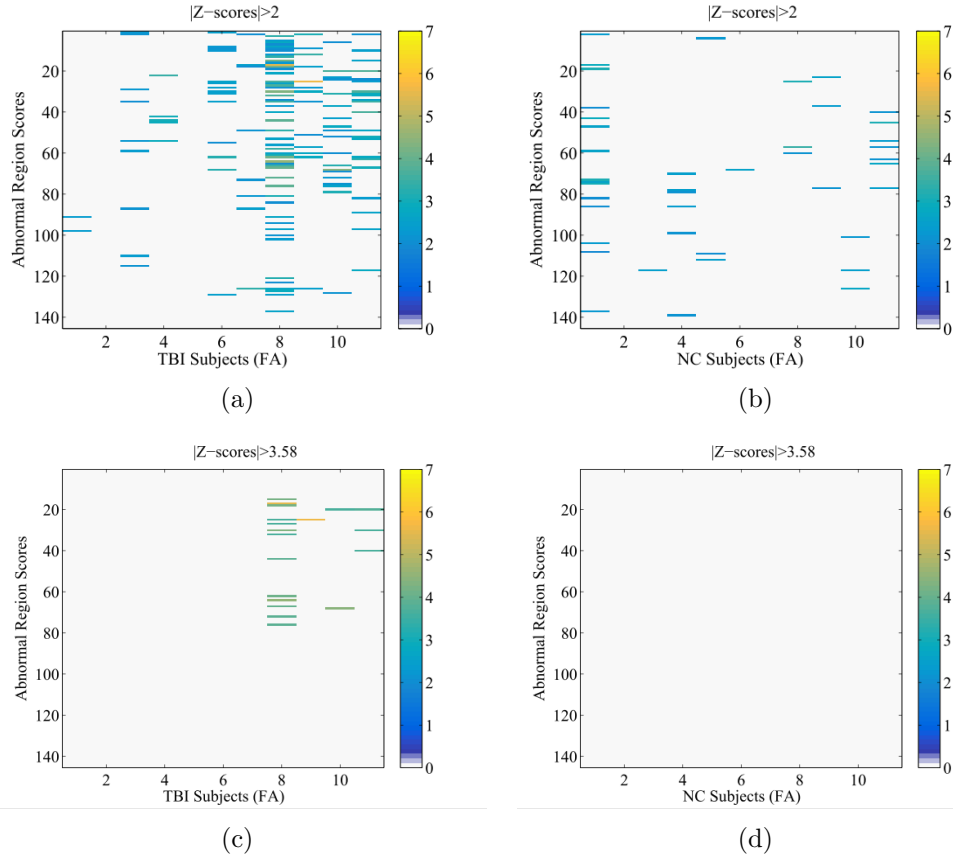


Figure 7: Abnormality maps based on z-scores for two different thresholds. Top: threshold is $|z| > 2$, Bottom: $|z| > 3.58$ (Bonferroni correction for the number of regions)